CRYSTAL STRUCTURE OF BACE AND USES THEREOF

[0001] This application claims the benefit of U.S. Provisional Application No. 60/234,576 filed September 22, 2000.

Field of the Invention

[0002] The present invention relates to the three dimensional crystal structure of Beta-site APP Cleaving Enzyme (BACE), and to the use of this structure in rational drug design methods to identify agents that may interact with active sites of BACE. Such agents may represent new therapeutics in the treatment and/or prevention of Alzheimer's Disease.

Background of the Invention

[0003] A characteristic pathology of Alzheimer's Disease is the build up of insoluble amyloid plaques in the brain. These proteinaceous plaques are composed of a 4KDa, 42 amino acid fragment of β-Amyloid Precursor Protein (APP) and is termed Amyloid β-peptide (Aβ). The mechanism of Aβ production is hence of critical importance in understanding the onset and progress of Alzheimer's Disease. It has been shown that Aβ is derived from the proteolytic cleavage of a larger protein, β-amyloid precursor protein (APP). Two enzymes are responsible for this cleavage; first, the enzyme β-secretase cleaves APP at residue 671 (770aa isoform of APP numbering) and then γ-secretase cleaves at residue 716. More recently, the novel transmembrane aspartic protease BACE has been identified as being β-secretase. This protein is now a significant target in a therapeutic approach to Alzheimer's Disease. In rare cases of Alzheimer's Disease that are hereditary (Familial Alzheimer's Disease (FAD)) the disease phenotype has been isolated to mutations in the β-Amyloid Precursor Protein. One particular cohort, the 'Swedish mutation', exhibits a double mutation at the β-secretase cleavage site.

[0004] Based upon the role of BACE in Alzheimer's Disease, the elucidation of the three-dimensional structure of BACE, as well as its site of binding with APP, would have important implications in the treatment and/or prevention of Alzheimer's Disease and similar diseases associated with the presence of insoluble amyloid plaques composed the 42 amino acid fragment of APP in the brain.

Summary of the Invention

[0005] The present invention provides a crystal of BACE complexed with an APP inhibitor peptide, as well as the three dimensional structure of BACE as derived by x-ray diffraction data of the BACE/APP inhibitor peptide crystal. Specifically, the three dimensional structure of BACE is defined by the structural coordinates shown in Figure 1, ± a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5Å. The structural coordinates of BACE are useful for a number of applications, including, but not limited to, the visualization, identification and characterization of various active sites of BACE, and the BACE/APP inhibitor peptide complex, including the APP binding site. The active site structures may then be used to design various agents which interact with BACE, as well as BACE complexed with an APP protein or peptide, or related molecules.

[0006] The present invention is also directed to an active site of an APP binding protein or peptide, and preferably the APP peptide binding site of BACE that is elucidated and derived from the three dimensional structure of BACE as defined by the relative structural coordinates set forth in Figure 1, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å.

[0007] In one embodiment of the present invention, the active site of the APP binding protein or peptide, preferably the APP peptide binding site of BACE, comprises the relative structural coordinates according to Figure 1 of residues SER71, GLY72, LEU91, ASP93, GLY95, SER96, VAL130, PRO131, TYR132, THR133, GLN134, ILE171, ILE179, ILE187, ALA188, ARG189, PRO190, TRP258, TYR259, ASP284, LYS285, ASP289, GLY291, THR292,

THR293, ASN294, ARG296 and ARG368, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å.

[0008] In another embodiment, the active site of the APP binding protein or peptide, preferably the APP peptide binding site of BACE, comprises the relative structural coordinates according to Figure 1 of residues LYS70, SER71, GLY72, GLN73, GLY74, TYR75, LEU91, VAL92, ASP93, THR94, GLY95, SER96, SER97, ASN98, TYR129, VAL130, PRO131, TYR132, THR133, GLN134, GLY135, LYS136, TRP137, LYS168, PHE169, PHE170, ILE171, ASN172, SER174, TRP176, GLY178, ILE179, LEU180, GLY181, ALA183, TYR184, ALA185, GLU186, ILE187, ALA188, ARG189, PRO190, ASP191, ASP192, ARG256, TRP258, TYR259, TYR283, ASP284, LYS285, SER286, ILE287, VAL288, ASP289, SER290, GLY291, THR292, THR293, ASN294, LEU295, ARG296, GLY325, GLU326, ARG368, VAL370, LYS382, PHE383, ALA384, ILE385, SER386, GLN387, SER388, SER389, THR390, GLY391, THR392, VAL393, GLY395, ALA396 and ILE447, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å.

[0009] The present invention further provides a method for identifying an agent that interacts with an active site of BACE. The method comprises the steps of: (a) determining a putative active site of BACE from a three dimensional model of BACE using the relative structural coordinates of Figure 1, ± a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5Å; and (b) performing various computer fitting analyses to identify an agent which interacts with the putative active site.

[0010] The present invention also provides method for identifying an agent that interacts with an active site of an APP binding protein or peptide, preferably BACE. The method comprises the steps of: (a) generating a three dimensional model of an active site of an APP binding protein or peptide using the relative structural coordinates according to Figure 1 of residues SER71, GLY72, LEU91, ASP93, GLY95, SER96, VAL130, PRO131, TYR132, THR133, GLN134, ILE171, ILE179, ILE187, ALA188, ARG189, PRO190, TRP258, TYR259, ASP284, LYS285, ASP289, GLY291, THR292, THR293, ASN294,

ARG296 and ARG368, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å; and (b) designing an agent using the three dimensional model generated in step (a).

[0011] The present invention also provides another method for identifying an agent that interacts with an active site of an APP binding protein or peptide, preferably BACE. The method comprises the steps of: (a) generating a three dimensional model of an active site of an APP binding protein or peptide using the relative structural coordinates according to Figure 1 of residues LYS70, SER71, GLY72, GLN73, GLY74, TYR75, LEU91, VAL92, ASP93, THR94, GLY95, SER96, SER97, ASN98, TYR129, VAL130, PRO131, TYR132, THR133, GLN134, GLY135, LYS136, TRP137, LYS168, PHE169, PHE170, ILE171, ASN172, SER174, TRP176, GLY178, ILE179, LEU180, GLY181, ALA183, TYR184, ALA185, GLU186, ILE187, ALA188, ARG189, PRO190, ASP191, ASP192, ARG256, TRP258, TYR259, TYR283, ASP284, LYS285, SER286, ILE287, VAL288, ASP289, SER290, GLY291, THR292, THR293, ASN294, LEU295, ARG296, GLY325, GLU326, ARG368, VAL370, LYS382, PHE383, ALA384, ILE385, SER386, GLN387, SER388, SER389, THR390, GLY391, THR392, VAL393, GLY395, ALA396 and ILE447, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å; and (b) designing an agent using the three dimensional model generated in step (a).

[0012] Finally, the present invention provides agents, and preferably inhibitors, identified using the foregoing methods. Small molecules or other agents which inhibit or otherwise interfere with the ability of BACE to cleave APP may be useful in the treatment and/or prevention of Alzheimer's Disease.

[0013] Additional objects of the present invention will be apparent from the description which follows.

Brief Description of the Figure

[0014] Figure 1 provides the atomic structural coordinates for BACE and the APP inhibitor peptide as derived by X-ray diffraction of a crystal of the BACE

and APP inhibitor peptide complex. "Atom type" refers to the atom whose coordinates are being measured. "Residue" refers to the type of residue of which each measured atom is a part - i.e., amino acid, cofactor, ligand or solvent. The "x, y and z" coordinates indicate the Cartesian coordinates of each measured atom's location in the unit cell (Å). "Occ" indicates the occupancy factor. "B" indicates the "B-value", which is a measure of how mobile the atom is in the atomic structure ($Å^2$).

Detailed Description of the Invention

[0015] As used herein, the following terms and phrases shall have the meanings set forth below:

[0016] Unless otherwise noted, "BACE" is Beta-site APP Cleaving Enzyme, and is the β-secretase enzyme that cleaves β-amyloid precursor protein (APP) at residue 671 (770aa isoform of APP numbering). After cleavage of APP by BACE, the remaining APP is cleaved at residue 716 by γ-secretase, leaving a 42 amino acid fragment of APP that is found in the proteinaceous plaques of Alzheimer's patients. The amino acid sequence of BACE preferably has the amino acid sequence deposited with Swiss Prot under accession number P56817, including conservative substitutions. As used herein, BACE also includes "BACE peptides," which are molecules having less than the complete amino acid sequence of BACE. Preferably, BACE peptides include the active site in which BACE binds to and cleaves APP. Most preferably, the BACE peptide corresponds to amino acid residues 58-447 set forth in Figure 1 ("$BACE_{58-447}$"), including conservative substitutions.

[0017] "APP" is β-amyloid precursor protein having the amino acid sequence deposited with Swiss Prot under accession number CAA31830, including conservative substitutions. As used herein, APP also includes "APP peptides," which are molecules having less than the complete amino acid sequence of APP. Preferably, APP peptides include the active site in which APP is cleaved by BACE.

[0018] An "APP inhibitor peptide" is a peptide which inhibits binding between BACE and APP. Preferably, the APP peptide has the amino acid sequence SER-GLU-VAL-ASN-Sta-VAL-ALA-GLU-PHE, where Sta is rare amino acid (S)-Statine.

[0019] An "APP binding protein or peptide" is a protein or peptide that binds APP and has a APP binding site, and includes but is not limited to BACE and BACE peptides.

[0020] Unless otherwise indicated, "protein" shall include a protein, protein domain, polypeptide or peptide.

[0021] "Structural coordinates" are the Cartesian coordinates corresponding to an atom's spatial relationship to other atoms in a molecule or molecular complex. Structural coordinates may be obtained using x-ray crystallography techniques or NMR techniques, or may be derived using molecular replacement analysis or homology modeling. Various software programs allow for the graphical representation of a set of structural coordinates to obtain a three dimensional representation of a molecule or molecular complex. The structural coordinates of the present invention may be modified from the original set provided in Figure 1 by mathematical manipulation, such as by inversion or integer additions or subtractions. As such, it is recognized that the structural coordinates of the present invention are relative, and are in no way specifically limited by the actual x, y, z coordinates of Figure 1.

[0022] An "agent" shall include a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound or drug.

[0023] "Root mean square deviation" is the square root of the arithmetic mean of the squares of the deviations from the mean, and is a way of expressing deviation or variation from the structural coordinates of BACE described herein. The present invention includes all embodiments comprising conservative substitutions of the noted amino acid residues resulting in same structural coordinates within the stated root mean square deviation.

[0024] The numbering of the amino acid residues identified in Figure 1 are based on the numbering of the full length BACE protein from the start of the signal sequence. It will be obvious to the skilled practitioner that the numbering of the amino acid residues of BACE may be different than that set forth herein or may contain certain conservative amino acid substitutions that yield the same three dimensional structures as those defined in Figure 1. Corresponding amino acids and conservative substitutions in other isoforms or analogues are easily identified by visual inspection of the relevant amino acid sequences or by using commercially available homology software programs (e.g., MODELLAR, MSI, San Diego, CA).

[0025] "Conservative substitutions" are those amino acid substitutions which are functionally equivalent to the substituted amino acid residue, either by way of having similar polarity, steric arrangement, or by belonging to the same class as the substituted residue (e.g., hydrophobic, acidic or basic) and includes substitutions having an inconsequential effect on the three dimensional structure of BACE, with respect to the use of this structure for the identification and design of agents which interact with BACE, for molecular replacement analyses and/or for homology modeling.

[0026] As used herein, an "active site" refers to a region of a molecule or molecular complex that, as a result of its shape and charge potential, favorably interacts or associates with another agent (including, without limitation, a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound, antibiotic or drug) via various covalent and/or non-covalent binding forces. Preferably, the active site of BACE corresponds to the site in which BACE cleaves the APP molecule.

[0027] As such, the active site of BACE may include, for example, both the actual site in which BACE binds and cleaves APP, as well as accessory binding sites adjacent or proximal to the actual binding site that nonetheless may affect the ability of BACE to bind and cleave APP, either by direct interference with the actual site of binding or by indirectly affecting the steric conformation or charge potential of the BACE molecule and thereby preventing

or reducing the ability of BACE to bind to APP at the actual binding site. As used herein, an active site also includes BACE or BACE analog residues which exhibit observable NMR perturbations in the presence of a binding ligand, such as APP or an APP peptide. While such residues exhibiting observable NMR perturbations may not necessarily be in direct contact with or immediately proximate to ligand binding residues, they may be critical to BACE residues for rational drug design protocols.

[0028] The present invention is directed to a crystallized complex of BACE and an APP inhibitor peptide that effectively diffracts X-rays for the determination of the structural coordinates of the complex. As used herein, BACE preferably corresponds to $BACE_{58-447}$ as set forth in Figure 1, with the N-terminal domain consisting of amino acid residues 58-207 shown in Figure 1, and the C-terminal domain consisting of amino acid residues 208-447 shown in Figure 1. The APP inhibitor peptide is preferably SER-GLU-VAL-ASN-Sta-VAL-ALA-GLU-PHE.

[0029] Using the crystal complex of the present invention, X-ray diffraction data can be collected by a variety of means in order to obtain the atomic coordinates of the crystallized molecule or molecular complex. With the aid of specifically designed computer software, such crystallographic data can be used to generate a three dimensional structure of the molecule or molecular complex. Various methods used to generate and refine the three dimensional structure of a crystallized molecule or molecular structure are well known to those skilled in the art, and include, without limitation, multiwavelength anomalous dispersion (MAD), multiple isomorphous replacement, reciprocal space solvent flattening, molecular replacement, and single isomorphous replacement with anomalous scattering (SIRAS).

[0030] Accordingly, the present invention also provides the three dimensional structure of BACE as derived by x-ray diffraction data of the BACE/APP inhibitor peptide crystal. Specifically, the three dimensional structure of BACE is defined by the structural coordinates shown in Figure 1, ± a root mean square deviation from the backbone atoms of the amino acids of

not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. The structural coordinates of BACE are useful for a number of applications, including, but not limited to, the visualization, identification and characterization of various active sites of BACE, and the BACE/APP inhibitor peptide complex, including the APP or APP peptide binding site. The active site structures may then be used to design agents with interact with BACE, as well as BACE complexed with APP, an APP peptide or related molecules.

[0031] The present invention is also directed to an active site of an APP binding protein or peptide, preferably the APP peptide binding site of BACE, which comprises the relative structural coordinates according to Figure 1 of residues SER71, GLY72, LEU91, ASP93, GLY95, SER96, VAL130, PRO131, TYR132, THR133, GLN134, ILE171, ILE179, ILE187, ALA188, ARG189, PRO190, TRP258, TYR259, ASP284, LYS285, ASP289, GLY291, THR292, THR293, ASN294, ARG296 and ARG368, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, more preferably not more than 1.0Å, and most preferably not more than 0.5Å.

[0032] In another preferred embodiment, the active site of an APP binding protein or peptide, preferably the APP peptide binding site of BACE, comprises the relative structural coordinates according to Figure 1 of residues LYS70, SER71, GLY72, GLN73, GLY74, TYR75, LEU91, VAL92, ASP93, THR94, GLY95, SER96, SER97, ASN98, TYR129, VAL130, PRO131, TYR132, THR133, GLN134, GLY135, LYS136, TRP137, LYS168, PHE169, PHE170, ILE171, ASN172, SER174, TRP176, GLY178, ILE179, LEU180, GLY181, ALA183, TYR184, ALA185, GLU186, ILE187, ALA188, ARG189, PRO190, ASP191, ASP192, ARG256, TRP258, TYR259, TYR283, ASP284, LYS285, SER286, ILE287, VAL288, ASP289, SER290, GLY291, THR292, THR293, ASN294, LEU295, ARG296, GLY325, GLU326, ARG368, VAL370, LYS382, PHE383, ALA384, ILE385, SER386, GLN387, SER388, SER389, THR390, GLY391, THR392, VAL393, GLY395, ALA396 and ILE447, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, more preferably not more than 1.0Å, and most preferably not more than 0.5Å.

[0033] Another aspect of the present invention is directed to a method for identifying an agent that interacts with an active site of BACE comprising the steps of: (a) determining an active site of BACE from a three dimensional model of BACE using the relative structural coordinates of Figure 1, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, more preferably not more than 1.0Å, and most preferably not more than 0.5Å; and (b) performing computer fitting analysis to identify an agent which interacts with said active site. Computer fitting analyses utilize various computer software programs that evaluate the "fit" between the putative active site and the identified agent, by (a) generating a three dimensional model of the putative active site of a molecule or molecular complex using homology modeling or the atomic structural coordinates of the active site, and (b) determining the degree of association between the putative active site and the identified agent. Three dimensional models of the putative active site may be generated using any one of a number of methods known in the art, and include, but are not limited to, homology modeling as well as computer analysis of raw data generated using crystallographic or spectroscopy data. Computer programs used to generate such three dimensional models and/or perform the necessary fitting analyses include, but are not limited to: GRID (Oxford University, Oxford, UK), MCSS (Molecular Simulations, San Diego, CA), AUTODOCK (Scripps Research Institute, La Jolla, CA), DOCK (University of California, San Francisco, CA), Flo99 (Thistlesoft, Morris Township, NJ), Ludi (Molecular Simulations, San Diego, CA), QUANTA (Molecular Simulations, San Diego, CA), Insight (Molecular Simulations, San Diego, CA), SYBYL (TRIPOS, Inc., St. Louis. MO) and LEAPFROG (TRIPOS, Inc., St. Louis, MO).

[0034] The present invention also provides a method for identifying an agent that interacts with an active site of an APP binding protein or peptide, and preferably the APP peptide binding site on BACE. The method comprises the steps of: (a) generating a three dimensional model of an active site of an APP binding protein or peptide using the relative structural coordinates according to Figure 1 of residues SER71, GLY72, LEU91, ASP93, GLY95, SER96, VAL130,

PRO131, TYR132, THR133, GLN134, ILE171, ILE179, ILE187, ALA188, ARG189, PRO190, TRP258, TYR259, ASP284, LYS285, ASP289, GLY291, THR292, THR293, ASN294, ARG296 and ARG368, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, more preferably not more than 1.0Å, and most preferably not more than 0.5Å; and (b) designing an agent using the three dimensional model generated in step (a). In another preferred embodiment, the active site of the APP binding protein or peptide is generated using the three dimensional model defined by the relative structural coordinates according to Figure 1 of residues LYS70, SER71, GLY72, GLN73, GLY74, TYR75, LEU91, VAL92, ASP93, THR94, GLY95, SER96, SER97, ASN98, TYR129, VAL130, PRO131, TYR132, THR133, GLN134, GLY135, LYS136, TRP137, LYS168, PHE169, PHE170, ILE171, ASN172, SER174, TRP176, GLY178, ILE179, LEU180, GLY181, ALA183, TYR184, ALA185, GLU186, ILE187, ALA188, ARG189, PRO190, ASP191, ASP192, ARG256, TRP258, TYR259, TYR283, ASP284, LYS285, SER286, ILE287, VAL288, ASP289, SER290, GLY291, THR292, THR293, ASN294, LEU295, ARG296, GLY325, GLU326, ARG368, VAL370, LYS382, PHE383, ALA384, ILE385, SER386, GLN387, SER388, SER389, THR390, GLY391, THR392, VAL393, GLY395, ALA396 and ILE447, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, more preferably not more than 1.0Å, and most preferably not more than 0.5Å.

[0035] The effect of such an agent identified by computer fitting analyses on the APP binding protein or peptide may be further evaluated by obtaining or synthesizing the agent, and contacting the identified agent with the APP binding protein or peptide in order to determine the effect the agent has on the APP binding protein or peptide. Preferably, the APP binding protein or peptide is BACE (or a BACE peptide), and the agent is a potential inhibitor of binding between BACE (or a BACE peptide) and APP (or an APP peptide). Therefore, in the preferred embodiment, the agent is contacted with BACE (or a BACE peptide) in the presence of APP (or a APP peptide), to determine the ability of the agent to inhibit binding between BACE (or the BACE peptide) and APP (or

the APP peptide). Depending upon the action of the agent on the active site, the agent may act either as an inhibitor or activator of the BACE/APP binding. Assays may be performed and the results analyzed to determine whether the agent is an inhibitor (i.e., the agent may reduce or prevent binding affinity between BACE and APP), an activator (i.e., the agent may increase binding affinity between BACE and APP), or has no effect on the interaction between BACE and APP. Agents identified using the foregoing methods, and preferably inhibitors of BACE cleavage of APP, may then be tested as therapeutics in the treatment and/or prevention of Alzheimer's Disease, and other diseases that are also characterized by the presence of the 42 amino acid fragment of APP in the proteinaceous plaques of the brain.

[0036] Various molecular analysis and rational drug design techniques are further disclosed in U.S. Patent Nos. 5,834,228, 5,939,528 and 5,865,116, as well as in PCT Application No. PCT/US98/16879, published WO 99/09148, the contents of which are hereby incorporated by reference.

[0037] Finally, the present invention is also directed to the agents, and preferably the inhibitors, identified using the foregoing methods. Such agents may be a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound, or drug, and preferably are small molecules that effectively inhibit binding between BACE and APP or an APP peptide. Such molecules may be useful in treating, preventing or inhibiting progression of Alzheimer's Disease.

[0038] The present invention may be better understood by reference to the following non-limiting Example. The following Example is presented in order to more fully illustrate the preferred embodiments of the invention, and should in no way be construed as limiting the scope of the present invention.

Example 1

A. Methods

[0039] Cloning of Human BACE1. Human polyA+ mRNA from whole brain (Clontech) was converted to cDNA by random-priming using Thermoscript

RT-PCR System, according to the manufacturer's protocol (Lifetechnologies). This cDNA was amplified by PCR using the forward and reverse primers, 5' GCTCTAGAACCCAGC ACGGCATCCGGCTG 3' (XbaI site indicated by underlined sequence; nts. 517-537 in accession no. AF190725) and 5' CCAAGCATGCGGCCGCAATAGGCTATGGTCA TGAGGGTTGAC 3' (NotI site indicated by underlined sequence; nts. 1809-1833; bold "A" indicates additional nucleotide to permit in-frame translation of the Fc chimera; see below), respectively. PCR was accomplished using Expand Long Polymerase kit according to the manufacturer's conditions (Roche Biochemicals; buffer #3), with PCR cycling consisting of an initial denaturing step at 95°C for 3min, 30-40 cycles of denaturation at 94°C for 30sec, annealing at 65°C for 30sec, elongation at 68°C for 1min 30sec, followed by a final elongation at 68°C for 5min. The PCR products were run on a 1% agarose gel. The appropriate band was cut out of the gel, purified by Quantum Prep Freeze 'N Squeeze DNA Extraction Columns (Bio-Rad), and cloned into the SpeI/NotI sites of the mammalian expression vector, pED/Fc (Kaufman, RJ et al., 1991, Nucl. Acids. Res. 19:4485-4490).

[0040] An intermediate construct contained the honey bee meletin secretory leader fused to the the prodomain and coding region of BACE1, just upstream to the predicted transmembrane domain of BACE1 (Vassar, R. et al., 1999, Science 286:735-741). The absence of the predicted hydrophobic transmembrane domain in this construct would permit soluble secreted BACE.Fc protein to be extracted from the conditioned medium. Downstream of BACE1 was an engineered enterokinase cleavage site followed by sequence encoding the Fc portion of immunoglobulin IgG. The final construct contained the BACE1.Fc gene, flanked by SalI and EcoRI in pED/Fc, cloned into the SalI/EcoRI sites of the mammalian expression vector, pHTop, a derivative of pED, in which the majority of the adenovirus major late promoter was replaced by six repeats of a bacterial tetracycline operator (described in Gossen et al, 1992, PNAS, 89:5547-5551). Sequencing of the BACE1.Fc recombinant gene was

accomplished by BigDye terminator dideoxy sequencing using an ABI3700. Sequence analyses was accomplished using DNAstar software package.

[0041] Expression of Human BACE1. The vector, pHTOP, with the BACE1.Fc insert, contains the dihydrofolate reductase gene and when introduced in the cell line CHO/A2 (see description below) functions very efficiently and high expressers can be selected by isolating cells surviving in high methotrexate concentrations.The CHO/A2 cell line is derived from CHO DUKX B11 (Urlaub and Chasin, 1980, PNAS USA 77:4216-4220) by stably integrating a transcriptional activator (tTA), a fusion protein between the Tet repressor and the herpes virus VP16 transcriptional domain (Gossen et al). A CHO cell line expressing extracellular BACE1.Fc was established by transfecting (lipofection) pHTopBACE1.Fc into CHO/A2 cells and selecting clones in 0.02 and 0.05 μM methotrexate. The conditioned media from multiple clones were screened by Western blot using a (mouse) anti-human IgG.Fc HRP antibody. The same clones were also metabolically labeled with 35 S (met/cys). The best clone, determined by virtue of its high expression, was one which resulted from 0.05 μM MTX selection and was chosen to be scaled up for roller bottle conditioned media production (4 Liters). The conditioned media was then used for purification. The expressed protein has residues 22-460 and nine extra residues at the C-terminal (an artefact from cloning and remains of the EK cleavage site).

[0042] Purification of BACE1. For the purification of BACE the 102 liters of conditioned media was used. During purification the activity of the enzyme was estimated at room temperature by continuously monitoring the fluorescent intensity for 5-10 min. at 420 nm (ext – 320 nm) Abz-Ser-Glu-Val-Asn-Leu-Asp-Ala-Glu-Phe-Arg-Dpa (Abz = Amino benzoic acid, Dpa = 9,10-diphenylanthracene) as the substrate. The reaction mixture contained 20 μM of substrate, different amounts of enzyme in 0.5 ml of 20 mM Tris-HCl pH 8.0 and 100 mM NaCl. The concentrated material of conditioned media(1.6 l) was applied to column (2.8 x 12 cm) containing ImmunoPure Immobilized Protein A

agarose (Pierce, Il, USA) equilibrated in PBS buffer. The speed of application was 2 ml/min. The column was washed with 1 litre of PBS buffer and the BACE-Fc protein was eluted by ImmunoPure IgG Elution Buffer (Pierce, Il, USA). The fractions containing protein were immediately neutralized by 1 M Tris-HCl to pH 8.0.

[0043] The obtained protein material was treated with Enterokinase at 25°C. The ratio of BACE-Fc to Enterokinase was 3000:1 and the time of reaction was 3 hrs. The reaction was stopped by removing Enterokinase from reaction mixture by applying the protein to a column (1 x 5 cm) containing soybean trypsin inhibitor agarose (Sigma, Mo, USA) equilibrated in 20 mM Tris-HCL pH 8.0 containing 100 mM NaCl (speed was 1 ml/min). The flow through material contained BACE and cleaved Fc. Cleaved Fc was removed from BACE by flowing through a protein A column equilibrated in 20 mM Tris-HCl pH 8.0.

[0044] BACE was partially de-glycosylated using PNGase F (New England Labs., Ma, USA). 8-9 μg of PNGase was added to 1 mg of BACE and the incubation was carried at 37° C for 16 hrs. The additional 5-6 μg of PNGase was added to each mg of BACE and incubation was continued for another 4 hrs. The purified BACE was separated from PNGase by HPLC size-exclusion chromatography using 21.5 x 30 cm G-3000SW column (TosoHaas, Pa, USA) equilibrated in 20 mM tris-HCL pH 8.0 containing 200 mM NaCl. (Speed of elution was 3 ml/min). The purified BACE was concentrated and used for crystallization experiments.

[0045] N-terminal sequencing of purified BACE reveals a mixture of protein species, with the major sample having the processing domain cleaved and beginning at residue 47 (all numbering refers to full length BACE;accession code: A59090) and a minor sample which had not been cleaved beginning at residue 22. A smaller sample with sequence MTIAY was also detected.

[0046] Crystallization. The crystals were grown using the hanging drop vapour diffusion method. The protein was concentrated to mg/ml in 20mM Tris pH 7.5, 200mM sodium chloride. Inhibitor peptide sequence is SEVNStaVAEF,

where Sta is the rare amino acid (*S*)-Statine. It was concentrated to 100mM in 100% DMSO and mixed with concentrated protein in a two-fold peptide excess to form the complex. 1 μl of complex was added to 1 μl of well solution containing 100mM Sodium Cacodylate pH6.5, 25% PEG8K, 300mM lithium sulphate. Plate-like crystal clusters grew within one week to dimensions of 200 μm x 400 μm x 75 μm. Single crystals were transferred to a stabilizing, cryoprotectant solution which contained the well solution plus 25% Glycerol for a brief, 10 second, soak and then frozen in liquid nitrogen. X-ray diffraction crystals had space group I222, and unit cell parameters a=86.627, b=130.861, c=130.729, and $\alpha = \beta = \gamma = 90°$.

B. Results

[0047] Structure Determination and Overall Fold. Full length BACE was expressed in CHO cells as a fc fusion protein and, after purification, cleavage and partial deglycosylation, complexed with peptide inhibitor and crystallized. Crystals diffracted to 2.3Å and the structure was solved using the technique of molecular replacement. The search model used was derived from cod atlantic Pepsin and contained 63% of the final number of atoms. The density modified maps obtained using a poly-alanine version of the search model (39% of the final atoms) provided sufficient information to build all but 12 amino acids. The final model contains residues from 59 to 448 (using full length numbering), all 9 residues of the statine inhibitor and 360 water molecules. Of the four predicted N-linked glycosylation sites only two have interpretable electron density.

[0048] The overall shape of the BACE protein is spherical and is composed of two distinct domains, an N-terminal (58-207) and a C-terminal (208-447). With the first thirteen amino acids (58-71) being packed against residues 238-243. There is a significant cleft-like channel across one surface of the interface between the domains. This contains the inhibitor peptide and conserved aspartic acid motifs that define the active sites of aspartic proteases.

[0049] The N-terminal domain is composed of a single a-helix preceeding the loop joining the two domains and thirteen β-strands. The larger C-terminal domain has a total of seventeen β-strands and three α-helices. The overall topology is characterised by an eight stranded antiparallel interdomain β-sheet. This central sheet comprises the majority of the active site residues including the two conserved aspartates (one from each domain:93 and 289). Asp93 and Asp289 define the position of a pseudo two-fold axis for the central β-sheet. Outside of this symmetry the two domains differ significantly. The N-terminal domain has an extra two strands extending the central sheet. In addition, there are two anti-parallel β-sheets above and below the central sheet composed of three and four β-strands respectively. Residues from the upper sheet (131-135) fold over the active site aspartates and form a 'flap' over the centre of the peptide binding cleft.

[0050] The C-terminal domain contains two lobes in addition to the strands which from the central β-sheet. These are weakly homologous to known aspartic protease structures. The binding pocket for the P1`and P3`positions are instead derived from three β-turns 388-391, 284-286 and 255-261.

[0051] There are a total of six cysteine residues in BACE. Each of these is involved in a disulphide interaction. The pattern of disulphide crosslinking, Cys278-Cys443, Cys380-Cys330, Cys420-Cys216 are unique in the aspartic proteases known to date.

[0052] A novel aspartic protease. The first attempts to study the relationship of function to structure of an Apartic proteases began in the 1930s with Pepsin. Since then this rich field of research has been successfully applied to the design of clinically used inhibitors in only one system; HIV protease. The reasons for this are related more to the validity of the pharmacological target than the efficacy of inhibitors. β-secretase has been described as a novel protease and has been shown to be linked to the onset and progression of Alzheimer's disease.

[0053] From a gross viewpoint the overall fold and domain organization is very similar to that of a canonical aspartic protease. The comparison at a more detailed level reveals a significant number of differences. The active site is characterized by two aspartic residues surrounded by a conserved set of hydrogen bonds termed a 'fireman's grip'. This is reproduced in the -secretase structure presented here. The characteristic flap which wraps over the active site in pepsin is absent from the C-terminal domain in a manner analogous to cathepsin D. In β-secretase the critical main chain amide hydrogen bond to the carboxyl group of statine is maintained by Thr133 from this flap. The amide of the statine makes a hydrogen bond to the carboxyl group of Gly95, emphasizing that the statine residue occupies both the P1 and P1` position.

[0054] Enzyme Mechanism. It has been shown that β-secretase cleavage is dependent on proximity to the cell membrane. Both β-secretase and its substrate APP have putative transmembrane regions. Our expressed BACE construct finishes one amino acid before the predicted transmembrane region. The final residue in the current structure is Ile447, thirteen residues away from the beginning of the putative transmembrane domain. In the current crystal structure Ile447 is only 6Å away from the P3 Glutamic acid of the inhibitor suggesting a role for the remaining C-terminal residues in the enzyme mechanism. The Statine residue of the inhibitor peptide is bound at the S1 position within the active site. The position of the C-3 hydroxyl group, coplanar to and within hydrogen bonding distance of both aspartate 93 and 289 carboxyl groups, confirms that the rare amino acid mimics the tetrahedral transition state i.e the intermediate of peptide-bond hydrolysis. The distance between the oxygen atoms of Asp93 and Asp289 is 2.8Å, strongly suggesting a shared proton atom and a classic aspartic protease pK profile for these side-chains and a common enzyme mechanism to other known aspartic proteases.

[0055] Inhibitor binding. The inhibitor peptide binds in an extended form along a 20Å groove formed at the interface between the domains. The

conserved catalytic aspartic residues lie at the middle of this groove. The bound peptide consists of 8 amino acids plus a statine amino acid at position 5. There is contiguous electron density for the whole peptide. The statine based inhibitor used in this study has been show to inhibit the β-secretase enzyme with nanomolar efficiency. The peptide sequence is based on the P10 to P4` APP751 Sweedish family mutation. This mutation of a Lys-Asn at the P2 position and Met-Leu at the P1 position is strongly correlated to the early onset of Alzheimer's disease. The inhibitor peptide utilizes Statine's Leucyl like side-chain to explore this interaction. Due to the di-peptide nature of Satine the P1` position of the substrate is shifted to P2` leaving an empty S1` pocket. The β-secretase enzyme appears to have a novel preference for an apartate or glutamate at the P1` position whereas other aspartic proteases show a preference for hydrophobic residues. This unusual preference for a negatively charged P1' amino acid is explained by the guanadinium group of Arg189 forming part of the putative S1` pocket. Even at the acidic pH optima of BACE the arginine side chain would form a positively charged environment for the possibly protonated carboxyl side-chain atoms.

[0056] The S1 and S3 binding pockets are a contiguous, hydrophobic pocket formed by the side-chain of residues Tyr132, Phe169, Ile171, Trp176, Ile179 and main chain atoms of Gly74, and Gln73. This packing of inhibitor P1 and P3 side chains has been seen in previous aspartic protease complexes.

[0057] The canonical APP cleavage site for b-secretase appears to have a preference for a small hydrophobic residues at the P2` position. The side chain of the valine residue bound in the putative S2` site of β-secretase appears to not make any significant interactions with the protein, its main chain however forms a tight set of hydrogen bonds to the backbone carboxyl of Gly 95 and the sidechain OH of Tyr259. In turn, Tyr259 is held rigidly in place by an edge-pi interaction with Trp258, which packs against the guanadinium group of Arg256.

[0058] Swedish mutation. Autosomal dominant mutations identified on the β-amyloid precusor protein have been correlated to early-onset cases of

Alzheimer's disease. These have been shown to cluster around the three canonical cleavage sites. A double (the so-called Swedish) mutation of Lys670-Met671 (770aa isoform of APP numbering) to Asn-Leu causes an increase in the overall quantity of Aβ detectable in the plasma and in the medium of cultured fibroblasts from carriers of the Swedish mutation. These two amino acids lie at the P2 and P1 positions of the β-secretase active site. The statine based inhibitor used here is based on this Swedish mutation. A methionine at position P1 would clearly be accommodated but would loose the van Der Waal's complentarity exhibited by the statine side-chain to Leu90 and Ile178. The Cϵ atom of the methione would make supplement the hydrophobic interaction to Phe169.

Table 1

Residues of BACE Within 4Å of Peptide Inhibitor

SER71, GLY72, LEU91, ASP93, GLY95, SER96, VAL130, PRO131, TYR132, THR133, GLN134, ILE171, ILE179, ILE187, ALA188, ARG189, PRO190, TRP258, TYR259, ASP284, LYS285, ASP289, GLY291, THR292, THR293, ASN294, ARG296, ARG368

Residues of BACE Within 8Å of Peptide Inhibitor

LYS70, SER71, GLY72, GLN73, GLY74, TYR75, LEU91, VAL92, ASP93, THR94, GLY95, SER96, SER97, ASN98, TYR129, VAL130, PRO131, TYR132, THR133, GLN134, GLY135, LYS136, TRP137, LYS168, PHE169, PHE170, ILE171, ASN172, SER174, TRP176, GLY178, ILE179, LEU180, GLY181, ALA183, TYR184, ALA185, GLU186, ILE187, ALA188, ARG189, PRO190, ASP191, ASP192, ARG256, TRP258, TYR259, TYR283, ASP284, LYS285, SER286, ILE287, VAL288, ASP289, SER290, GLY291, THR292, THR293, ASN294, LEU295, ARG296, GLY325, GLU326, ARG368, VAL370, LYS382, PHE383, ALA384, ILE385, SER386, GLN387, SER388, SER389, THR390, GLY391, THR392, VAL393, GLY395, ALA396, ILE447

[0059] All publications mentioned herein above, whether to issued patents, pending applications, published articles, deposited sequences, or otherwise, are hereby incorporated by reference in their entirety. While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of the disclosure that various changes in form and detail can be made without departing from the true scope of the invention in the appended claims.